# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 03-198133
(43)Date of publication of application : 29.08.1991

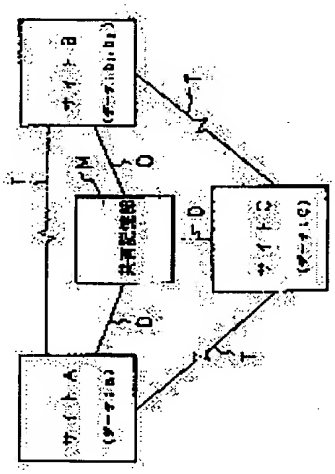
(51)Int.Cl. G06F 12/00
G06F 13/00

(21)Application number : 01-336655
(22)Date of filing : 27.12.1989

(71)Applicant : TOSHIBA CORP
(72)Inventor : MORIMOTO YOJIRO

## (54) DECENTRALIZED DATA BASE PROCESSING SYSTEM

(57)Abstract:
PURPOSE: To process at a high speed those many message that are produced in the decentralized data base processing by providing a shared storage means for plural computer sites in addition to a communication circuit which connects the computer sites to each other.
CONSTITUTION: A shared storage means M is provided for computer sites A-C in addition to a communication circuit T which connects the sites A-C to each other. Thus the data can be transferred among the sites A-C via a means M in parallel with the communication of data carried out via the circuit T. In other words, the means M that is newly provided among the sites A-C is used together with the conventional circuit T. Then the means M and the circuit T are properly used in accordance with their using conditions. Thus it is possible to eliminate the bottleneck of data communication, i.e., a problem produced so far in the decentralized data base processing. Then the data reference processing efficiency is improved.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

⑲日本国特許庁（JP）　⑪特許出願公開

⑫公開特許公報（A）　平3－198133

| ⑤1Int. Cl.⁵ | 識別記号 | 庁内整理番号 |
|---|---|---|
| G 06 F 12/00 | 301 S | 8944－5B |
| 13/00 | 351 J | 7459－5B |

⑬公開　平成3年(1991)8月29日

審査請求　未請求　請求項の数 11　(全15頁)

⑤4発明の名称　分散データベース処理方式

⑳特　願　平1－336655
㉒出　願　平1(1989)12月27日

⑦2発明者　森本　陽二郎　神奈川県川崎市幸区小向東芝町1番地　株式会社東芝総合研究所内
⑦1出願人　株式会社東芝　神奈川県川崎市幸区堀川町72番地
⑦4代理人　弁理士　鈴江　武彦　外3名

明　細　書

1．発明の名称

分散データベース処理方式

2．特許請求の範囲

(1)　データベースをそれぞれ備えた複数の計算機サイトと、これらの計算機サイトを相互に結合する通信回線からなり、上記各計算機サイト上にそれぞれ散在しているデータ群を取り扱う分散データベースシステムにおいて、

前記各計算機サイトを相互に結合してデータ通信に供される前記通信回線とは独立に前記各計算機サイトで共有される共有記憶手段を設け、この共有記憶手段を介して前記各計算機サイト間でのデータの受け渡しを行い得るようにしたことを特徴とする分散データベース処理方式。

(2)　通信回線を介するデータ通信と共有記憶手段を利用したデータの受け渡しとの切り替えは、通信要求のあるデータの重要度に応じて選択的に行われることを特徴とする請求項(1)に記載の分散データベース処理方式。

(3)　データベースをそれぞれ備えた複数の計算機サイトと、これらの計算機サイトを相互に結合する通信回線からなり、上記各計算機サイト上にそれぞれ散在しているデータ群を取り扱う分散データベースシステムにおいて、

前記各計算機サイトを相互に結合してデータ通信に供される前記通信回線とは独立に前記各計算機サイトで共有される共有記憶手段を設けると共に、

前記各計算機サイトに、前記通信回線の負荷を検出する手段と、この負荷状態検出結果に従って前記通信回線を介するデータ通信と前記共有記憶手段を介するデータの受け渡しとを切り替え制御する手段とをそれぞれ設けたことを特徴とする分散データベース処理方式。

(4)　各計算機サイト間での共有記憶手段を介するデータの受け渡しは、通信回線を介するデータ通信の負荷が高いときに行われることを特徴とする請求項(3)に記載の分散データベース処理方式。

(5)　通信回線を介するデータ通信の負荷は、通

信要求の待ち行列の長さから求められることを特徴とする請求項(4)に記載の分散データベース処理方式。

(6) 通信回線を介するデータ通信の負荷は、通信の履歴または単位時間当りの通信統計量から求められることを特徴とする請求項(4)に記載の分散データベース処理方式。

(7) データベースをそれぞれ備えた複数の計算機サイトと、これらの計算機サイトを相互に結合する通信回線からなり、上記各計算機サイト上にそれぞれ散在しているデータ群を取り扱う分散データベースシステムにおいて、

前記各計算機サイトを相互に結合してデータ通信に供される前記通信回線とは独立に前記各計算機サイトで共有される共有記憶手段を設けると共に、

前記各計算機サイトに、前記共有記憶手段の利用状況を検出する手段と、この共有記憶手段の利用状況に従って前記通信回線を介するデータ通信と前記共有記憶手段を介するデータの受け渡しとを切り替え制御する手段とをそれぞれ設けたことを特徴とする分散データベース処理方式。

(8) 各計算機サイト間での通信回線を介するデータ通信は、共有記憶手段の利用状況が高く、高負荷状況にあるときに行われることを特徴とする請求項(7)に記載の分散データベース処理方式。

(9) 共有記憶手段の利用状況は、通信要求の待ち行列の長さから求められることを特徴とする請求項(8)に記載の分散データベース処理方式。

(10) 共有記憶手段の利用状況は、通信の履歴または単位時間当りの通信統計量から求められることを特徴とする請求項(8)に記載の分散データベース処理方式。

(11) 共有記憶手段の利用状況は、共有記憶手段の空き領域の量から求められることを特徴とする請求項(8)に記載の分散データベース処理方式。

3. 発明の詳細な説明

[発明の目的]

(産業上の利用分野)

本発明は分散型のデータベース管理システムにおけるデータ参照要求に対する処理を効率的に行い得る分散データベース処理方式に関する。

(従来の技術)

近時、種々のデータベースを利用した情報処理システムが盛んに開発されている。特に最近ではデータベースの大規模化に伴い、複数の計算機サイトにデータベースを分散配置し、これらの計算機サイトを通信回線を介して相互に結合することで、上記各計算機サイトに分散配置されたデータベースを相互利用するようにした分散データベースシステムが開発されている。

即ち、この種の分散データベースシステムは、例えば第9図に示すようにデータベースを含む複数(ここでは3個)の計算機サイトA, B, Cを通信回線Tを介して相互に結合して構築される。これらの各計算機サイトは、それぞれ独立した計算機システムとして動作する機能を有し、例えば第10図に示すようにCPU1を主体とし、主記憶2を備えると共に、入出力制御装置3を介してデータベースを蓄積した外部記憶装置4に接続して構成される。そして上記CPU1の制御の下で動作し、主記憶2上に分散データベース管理部5と分散データベース通信制御部6とを構築し、通信回線管理部7に接続された通信回線Tを介して他の計算機サイトとの間で適宜データの受け渡しを行ってデータベース参照を行うものとなっている。

尚、データベース参照は、データベースに対する読み出し・書き込み・消去・追加・初期化等の計算機サイトにおける操作手続きを示すものである。

この種の分散データベースシステムは、物理的に分かれて構築される複数のデータベースを、複数の計算機サイトにて協調して利用することを目的として構築されるもので、既存の複数のデータベースを後で論理的に結びつけることによって、或いは予め決められた分散形態に従ってデータベースを分散させることにより実現される。いずれの場合にしろ、データベースは計算機サイト毎に分離されて構築される。

しかしてこのような分散データベースシステムによれば、個々のデータベースを相互に協調させることにより高度な情報を得ることが可能となり、また故障等による危険を分散させることができる等の利点がある。更にはデータベース参照の為の計算機サイトでの負荷を分散することができ、各データベース毎にその運用に適した管理を行うことが可能である等の優れた利点を有する。これ故、今後益々数多く構築され、その利用の度合いが増えると予想される。

ところで複数の計算機サイトに分散したデータベースを、利用者側から見てあたかも1個のデータベースのように見せる分散データベース管理システムでは、複数の計算機サイトにまたがるデータ参照要求の発生に対しては、そのデータ参照要求を各サイトに対する部分参照要求に分解し、それぞれのサイトに対する参照要求や、排他制御等を行う為の制御情報を送受信する必要がある。その上で、各計算機サイト間でのデータの受け渡しが行われる。

このようにして計算機サイト間で通信される情報はメッセージと称される。そしてこの通信回線Tを介して通信されるメッセージは、複数の計算機サイトでの並行制御の為のデータロック情報や、操作完了を表すコミットメント情報等を構成する通信制御に関するメッセージと、参照要求に応じてデータベースから取り出したデータや、データベースに格納すべきデータ等のメッセージの2種類からなる。

しかして一般に、前者のメッセージはデータ量自体は少ないがその交信するメッセージの数が多く、これに対して後者のメッセージはデータ量が多いがそのメッセージ数は前者に比べ少ないという性質がある。そして1回のデータ参照要求を処理するだけでも前記計算機サイト間で多くのメッセージ通信を行うことが必要となる。特に、データのコピーが各サイトに存在する場合には、データ更新時に各コピーデータの一貫性を維持する為により多くのメッセージ通信を行う必要が生じる。

ところが現状では、第9図に示すようにこれらのメッセージを単一の通信手段（通信回線T）を用いて通信することになるので、その通信路が混雑し、通信処理に多くの時間が掛かることが否めない。しかも分散環境下でのデータ参照処理においては、個々の計算機サイトにおける演算処理に比較して、その通信に多くの処理時間を必要とする。これ故、データ参照処理の高速化に際して、メッセージ通信に要する時間の損失が大きな問題となっている。

このような問題点に関しては、例えば

『JEPPREY D.ULLMAN 著　国井・大保 訳，
"データベース・システムの原理" pp.497-548
日本コンピュータ協会発行 （1985.5.25）』

等に詳しく紹介されている。

（発明が解決しようとする問題点）

このように従来の分散データベースシステムでは、データ参照要求に応じて複数の計算機サイト間で通信すべきメッセージの量やその通信回数が多いにも拘らず、複数の計算機サイト間を単一の通信手段（通信回線T）を介して結合しているだけなので、その通信回線の混雑化等に起因してメッセージ通信に要する時間が長く掛り、データ参照処理の高速化を図ることが非常に困難となっている等の不具合があった。

本発明はこのような事情を考慮してなされたもので、その目的とするところは、複数の計算機サイト間でのメッセージ通信の効率化を図ることを可能とし、データ参照処理の高速化・効率化を図るようにした実用性の高い分散データベース処理方式を提供することにある。

［発明の構成］

（課題を解決するための手段）

本発明はデータベースをそれぞれ備えた複数の計算機サイトと、これらの計算機サイトを相互に結合する通信回線からなり、上記各計算機サイト上にそれぞれ散在しているデータ群を取り扱う分散データベースシステムに係り、

前記計算機サイトを相互に結合する通信回線とは別個に、前記各計算機サイトで共有される共有記憶手段を設け、通信回線を介するデータ通信と

並行して、上記共有記憶手段を介して前記各計算機サイト間でのデータの受け渡しを行い得るようにしたことを特徴とするものである。

特に各計算機サイトに、前記通信回線の負荷を検出する手段、或いは前記共有記憶手段の利用状況（負荷）を検出する手段を設け、この負荷状態検出結果に従って、或いはデータの重要度等に応じて前記通信回線を介するデータ通信と前記共有記憶手段を介するデータの受け渡しとを切り替え制御するようにしたことを特徴とするものである。

即ち、本発明は複数の計算機サイト間に新たに配置した共有記憶手段を、従来の通信回線と併せて利用し、例えばこれらの使用状況に応じて上記通信回線と共有記憶手段とを使い分けることにより、従来の分散データベース処理で問題となっていたデータ通信のボトルネックを解消し、データ参照処理の効率向上を実現するようにしたことを特徴とするものである。

（作　用）

本発明によれば、複数の計算機サイトを相互に結合する通信回線と、上記各計算機サイトに共有される共有記憶手段とを使い分けて計算機サイト間でのメッセージ通信を行うので、従来のように通信回線の混雑化等に起因してメッセージ通信に要する時間が長く掛る等の不具合がなくなる。そして、例えば上記通信回線の負荷状況や共有記憶手段の使用状況に応じて、またデータの重要度に応じて通信回線と共有記憶手段とを使い分けるので、その通信効率を効果的に高めることができる等の効果が奏せられる。

（実施例）

以下、図面を参照して本発明の一実施例に係る分散データベース処理方式について説明する。

第１図は実施例方式を適用して構成される分散データベースシステムの概略構成図で、ここでは３つの計算機サイトＡ，Ｂ，Ｃを用いて構成されている例を示している。しかして各計算機サイトＡ，Ｂ，Ｃは、通信回線Ｔを介して相互に接続されている。更にここでは、各計算機サイトＡ，Ｂ，Ｃに共有される共有記憶部Ｍが設けられている。

この共有記憶部Ｍは半導体メモリ等を利用して構成されるもので、前記各計算機サイトＡ，Ｂ，Ｃとの間でバス等の情報伝送路Ｄを介して接続され、各計算機サイトＡ，Ｂ，Ｃからそれぞれ高速にデータ参照し得るように構成されている。

このようにして他の計算機サイトとの間で通信回線Ｔを介して接続されると共に、前記共有記憶部Ｍを共有してなる各計算機サイトは、概略的には他の計算機サイトとの間での通信を制御する手段として、前記通信回線Ｔを利用した通信を管理する通信回線管理部と、前記共有記憶部Ｍへのデータ参照を管理する共有記憶管理部とを備えて構成される。そして、例えば通信回線負荷検出手段と共有記憶負荷検出手段とを用いて前記通信回線Ｔと共有記憶部Ｍの混み具合いをそれぞれ検出し、これらの負荷状態検出結果に従って前記通信回線管理部または共有記憶管理部にメッセージ通信要求を出し、通信回線Ｔまたは共有記憶部Ｍを選択的に介して他の計算機サイトとの間でデータの受け渡しを行う如く構成される。

即ち、計算機サイトは、例えば第２図に示すように構成される。この計算機サイトは、それだけで計算機としての処理を行う機能を持ち、ＣＰＵ１，主記憶２を主体として構成される。このＣＰＵ１に、その入出力を管理する為のＩ／Ｏチャネル等の入出力制御装置３を介してデータベース等を格納する磁気ディスク装置等の外部記憶装置４が接続される。

複数の計算機サイトに分散されたデータベースを管理する為の分散データベース管理部５は、そのデータ配置管理や質問処理，データベース参照制御、耐障害処理等の一連したデータベース管理を行うもので、一般にはソフトウェア上で実現され、稼働時には前記主記憶２上に設けられる。この分散データベース管理部５は、ソフトウェアの実現上、自サイトのデータベースを管理する部分と分散したデータベースを大域的に管理する部分とに分けて構成されることが多いが、１個のソフトウェアシステムとして実装することも可能である。

この分散データベース管理部5での処理により、メッセージ通信要求が発生すると分散データベース通信制御部6に通信要求が渡され、要求に応じた通信が実行される。この分散データベース通信制御部6も一般的にはソフトウェアで実現され、前記主記憶2上に実装される。尚、この分散データベース通信制御部6を前記分散データベース管理部5の一部として実現することもできるが、ここでは別個の手段として実現されるものとして説明する。

一方、通信回線管理部7は、前記通信回線Tを介する他のサイトとの間でのメッセージ通信を司るもので、例えばオペレーティングシステムに組み込まれて実現されたり、或いはオペレーティングシステムの外郭ソフトウェアとして稼働するように構成される。この通信回線管理部7により前記通信回線Tを介するメッセージ通信が管理される。ちなみに従来の分散データベース処理では、データベース参照時に発生したメッセージ通信の全てがこの通信回線管理部7を用いて実現されている。

しかしてこの計算機サイトが持つ新しい機能として、前記共有記憶部Mを用いたメッセージの送信と受信とを管理する為の共有記憶管理部8が設けられている点にある。この共有記憶管理部8は、基本的には分散データベースシステムが稼働している時、基本的には分散データベース処理以外の処理で発生した通信処理を共有記憶部Mを用いて実行するものである。この共有記憶管理部8を介して分散データベース参照処理以外の処理で発生するメッセージの通信が行われることにより、そのメッセージ通信が高速に行われるようになっている。

そしてこのような共有記憶管理部8により前記共有記憶部Mおよび情報伝送路Dが管理され、前記分散データベース通信制御部6からの分散データベース参照処理に応じたメッセージ通信要求に従って前記共有記憶部Mおよび情報伝送路Dを介して他の計算機サイトとの間でのメッセージの送受信が行われる。

この際、前記共有記憶部Mの管理については、予めその記憶領域の使い方等の規約を設けておいて各計算機サイトで共同で管理するようにしてもよいが、共有記憶部MにＣＰＵや管理プログラムを持たせておき、共有記憶部M自体にその管理手段を持たせておくことも可能である。勿論、この共有記憶部Mに計算機の機能を持たせておくことも可能であり、このようにすれば共有記憶部Mの管理をより簡単に実現することができる。

尚、ここでは共有記憶管理部8を分散データベース通信制御部6に接続した構成となっているが、物理的には他の手段と接続されていてもよい。即ち、分散データベース処理を行っていないときや分散データベース処理を行っていても高速通信を必要としないときは、例えば応用プログラム等で前記共有記憶部Mを使用することが可能である。つまり共有記憶部Mも計算機資源として活用可能であることからオペレーティングシステムの管理下に置く方が利用者にとって好都合である。従って、例えば共有記憶部Mの管理や参照処理の基本部分をオペレーティングシステムの内部に基本ルーチンとして組み込んでおき、これらを前記分散データベース通信管理部6や前記共有記憶管理部8で呼び出すような形態で実装しておくことも可能である。

このようにこの実施例に係る計算機サイトは、通信回線Tを介する計算機サイト間でのデータ通信を管理する為の通信回線管理部7に加えて、共有記憶部Mとの間でデータの受け渡しを行うことで該共有記憶部Mを介する他の計算機サイトとの間でのデータ通信を管理する為の共有記憶管理部8を備えたことを特徴としている。そして前記分散データベース通信制御部6の制御の下でこれらの通信回線管理部7と共有記憶管理部8とを選択的に起動し、前記通信回線Tまたは共有記憶部Mを介して他の計算機サイトとの間でデータ通信を行うように構成されていることを特徴としている。

このような特徴的な機能に加えて、この実施例に係る計算機サイトが備えている機能として、第2図に示すように通信回線負荷検出手段9と共有

記憶負荷検出手段10とがある。通信回線負荷検出手段9は前記通信回線管理部7での前記通信回線Tを介するメッセージ通信の混み具合いを検出する為の手段である。また共有記憶負荷検出手段10は前記共有記憶管理部8を用いた共有記憶部Mとの間でのメッセージ通信の混み具合いを検出する為の手段である。前記分散データベース通信制御部6はこれらの検出手段9,10から前記通信回線Tの混み具合いと，共有記憶部Mおよび情報伝送路Dの混み具合いとをそれぞれ求め、例えば以下に示すようにそのメッセージ通信を制御するものとなっている。

尚、通信回線負荷検出手段9と共有記憶負荷検出手段10におけるメッセージ通信の混み具合の検出は、例えば次のような判断基準の下で行われる。

① それぞれの通信手段に対する通信要求の待行列の長さを負荷の基準とする。即ち、通信要求の待行列の長さが長いほど通信が混んでいると判断する。

② 通信要求の待行列を形成する各要求要素の転送データ量を合計し、この通信待ちのデータ量を通信負荷の基準とする。即ち、通信待ちのデータ量の多い方が混んでいるとして判断する。

③ 通信の履歴や単位時間当りの通信回数，通信データ量等の統計量によりその通信負荷の程度を求める。具体的には、例えば最近よく使われている（使用頻度の高い）通信手段をその通信負荷が大きいと看做す。

④ 共有記憶部Mにおいて受信済みのメッセージを消去することを前提とした場合、共有記憶部Mに残されている空き領域（新たに通信に使用することができる領域）の量を共有記憶部Mの混み具合いの基準とする。この場合、空き領域が少ないほど、その通信負荷が混んでいることになる。

このような判断基準を選択的に用い、或いは適宜組み合わせて用いることで前記各通信手段の混み具合がそれぞれ判定される。

次にこのようにシステム構成された分散データベースシステムにおける処理手続きの例について説明する。

尚、ここではその説明を具体的にするべく、例えば端末機器などから計算機サイトAに次のようなデータ参照要求、或いはデータ参照要求を発生する源となる情報が届いたものとする。

『計算機サイトBのデータベースに格納されているデータb1の値から、計算機サイトAのデータベースに格納されているデータaの値を差し引き、その結果を計算機サイトCのデータベースに格納されているデータcの値から引き去る。そしてその結果を前記計算機サイトBのデータベースに新たなデータb2として書き込む。』

このときのデータ参照要求は、

［b2 ← c ← c－（b1－a）］

として表すことができる。

尚、［←］は値の代入を表し、式は右側ほど優先順位が高い。

このようなデータ参照要求の手続きを計算機サイトAを中心として処理する場合、そのデータ参照および演算手順は、例えば第3図のように表現される。この第3図における各命令および式の意味は、［／＊ ～ ＊／］により示され、また［A：a］は計算機サイトAのデータ項目aを表している。

しかしてこのデータ参照処理の例では、データ項目b1の値とデータ項目aの値との差を計算機サイトCに送り、この計算機サイトCにて新しいデータ項目cの値を計算して求め、これをそのデータベースに書き込む。そしてその結果を計算機サイトBのデータ項目b2に書き込むことになる。しかしここでは計算機サイトBでのデータ項目b2への書き込みは、計算機サイトAでの処理手続きとは直接関係がないことから、ここでは第3図に示す処理手順に従ってその処理が進められるものとする。

さて一般的にデータベースの参照要求は、複数の計算機サイトにおいて同時に複数発生することが多い。この現象は大規模なデータベースシステムとなるほど顕著に現われる。そこでこれらのデータベースの参照要求を効率よく処理する為に、ここでは1つの参照要求に対し、それが完結する

までＣＰＵ処理時間を割り当てるのではなく、例えば複数の参照要求に対して時間を区切ってＣＰＵ処理時間を割り当てるものとなっている。

この際、各計算機サイトでのデータの一貫性を保つ為の並行制御が必要となることから、例えばデータロックを行った後、そのデータを参照する等の手法がとられる。この場合、その参照要求を直列的に実行したのと同じ結果を与える手法が従来より種々提唱されているが、ここでは最も一般的な２相ロックプロトコルを採用したものとして説明する。また説明を簡単にする為に、ここではデータロックはデータの読み出しと書き込みのいずれに対しても行われるものとする。

またデータベース処理では、計算機や外部記憶装置，通信機器等に故障が起こったときにもそのデータベースを保全することが必要である。特に分散データベースの場合、故障発生の対象となる機器が多い為、データの更新には特に慎重をきたす必要がある。そこでこの例では２相コミットプロトコルを採用し、コミット前にデータの更新に関与した全サイトのコミット準備完了を確認した上でデータ参照処理を実行するものとしてその説明を進める。

しかしてこれらの方式を採用した場合、前記計算機サイトＡを中心とした処理を行った場合のメッセージの流れは、例えば第４図に示すようになる。但し、処理の途中における故障や、他のデータ参照要求とデータのロックの衝突、またデッドロックやオペレータによるアボート等がなく、ここでは順調に処理が進められるものとする。またこの例では処理の始めに、参照対象となる全てのデータをロックするものとして規定する。

第４図においてメッセージ［６］，［８］，［９］，［１０］はデータを転送するためのメッセージであり、それ以外は制御に関するメッセージである。尚、メッセージ［１］，［５］，［６］についてはそれぞれ２個のメッセージに分けることもあるが、メッセージを分けた場合にはその分、その通信回数は多くなる。またデータａの参照に関してもロック（ＬＯＣＫ）やコミット（ＣＯＭＭＩＴ）に関する操作が必要となるが、ここでは計算機サイトＡを中心として処理を進める為、この例では特にメッセージを送受信する必要はない。またこの例ではデータのＵＮＬＯＣＫは、それぞれのサイトでデータを実際に更新した後に行っているものとする。

先ず第４図に示すメッセージはそれぞれ次のような意味を持つ。

ＬＯＣＫ（ａ）：ａという名のデータ項目をロックする。

ＬＯＣＫ－ＥＮＤ：ロック処理が完了。

ＲＥＡＤ（ａ）：ａという名のデータ項目を読み出す。

ＤＡＴＡ（ａ）：ａという名を持つデータ項目の内容を送る。

ＷＲＩＴＥ（ａ）：データをａという名のデータ項目に書き込む。

ＣＯＭＭＩＴ１：１相目のコミット要求。

ＣＯＭＭＩＴ２：２相目のコミット要求。

ＣＯＭＭＩＴ－ＯＫ：コミット準備完了、或いはコミット完了。

この第４図に示す処理手続きから明らかなように、この例では１８回のメッセージ通信を必要とすることが判る。ちなみに従来ではこれらのメッセージ通信の全てを通信回線Ｔを介して実行していた。またここで説明するデータ参照要求は、当然のことながら他の計算機サイトＢ，Ｃでも同時に発生することがある。従ってこのようなデータ参照要求は、大規模な情報システムでは１秒間に数十回も発生することが予想される。そしてその通信回数と転送データ量はデータ参照要求に比例して多くなる為、従来ではメッセージ通信が大きなボトルネックとなり、システム全体のスループットに大きな制約を与えていた。

これに対して本発明の実施例では、メッセージの全てを通信回線Ｔを介して行うことなく、前記共有記憶部Ｍをも利用して行われる。この共有記憶部Ｍは前述したように半導体等の高速参照可能な記憶媒体により実現されており、また計算機サ

イトとはメモリバス等で結合されていることから、前記共有記憶部Mを利用する場合には通信回線Tを介してメッセージを転送する場合に比較してはるかに高速にメッセージ転送し得るようになっている。

第5図は上述した第4図に示すメッセージ通信を実現する処理手続きの流れを示す図であり、この第5図を参照してこの実施例におけるデータベース参照処理手順について説明する。

先ず、端末機器などから与えられたデータ参照要求が分散データベース管理部5にて解析され、その処理手順が決定される（ステップa）。この解析処理の結果、例えば第4図に示すようなメッセージの交換が必要であることが求められる。分散データベース管理部5は、このような処理手順に従って命令，或いは操作を処理する（ステップb）。

しかして命令がメッセージ送信要求の場合には（ステップc）、先ず前記分散データベース管理部5から分散データベース通信制御部6にメッセージが渡される（ステップd）。すると分散データベース通信制御部6は、前記通信回線負荷検出手段9と共有記憶負荷検出手段10に対してそれぞれの負荷状況を調べる為の要求を出す（ステップe）。このような要求を受けて前記通信回線負荷検出手段9および共有記憶負荷検出手段10は、それぞれ通信回線管理部7と共有記憶管理部8の利用状況を調べ、その通信負荷状態をそれぞれ求める（ステップf）。この負荷状況の判断は前述したように通信要求の待ち行列の長さを調べる等して行われる。

このようにして前記通信回線負荷検出手段9および共有記憶負荷検出手段10にて前記通信回線管理部7および共有記憶管理部8をそれぞれ参照して求められる通信回線Tおよび共有記憶部Mの各負荷状況の情報が前記分散データベース通信制御部6にそれぞれ返される（ステップg）。分散データベース通信制御部6は、このようにして求められる前記通信回線Tおよび共有記憶部Mの各負荷状況の情報に従い、所定の負荷判断基準に基いてメッセージ通信に供する通信手段を決定する（ステップh）。

そして共有記憶部Mを用いてメッセージ通信を行うか否かを判断された場合（ステップi）、共有記憶部Mを用いてメッセージを相手側の計算機サイトに送り（ステップj）、共有記憶部Mを用いない場合には通信回線Tを介してメッセージを相手側の計算機サイトに送る（ステップk）。

この通信回線Tおよび共有記憶部Mの各負荷状況の情報に従うメッセージ通信に供する通信手段を決定、つまり通信回線負荷検出手段9および共有記憶負荷検出手段10によりそれぞれ求められた混み具合いの判断は、例えば次のようにして行われる。

① 通信回線Tと共有記憶部Mのうちで、それぞれの管理部7,8における待ち行列の短い方を空いていると判断する。

② 双方の通信手段について、全ての転送待ちデータ量の合計の少ない方を空いていると判断する。

③ 通信回線Tおよび共有記憶部Mの各処理速度に基づき、現在の通信待ち行列の通信要求の全てが実行されるまでの時間を予測し、予測された時間の短い方を空いていると看做す。

④ いずれにも待ち行列がなく、双方とも空いている状態ならば通信回線Tを選択する。

④′ 逆にいずれも待ち行列がなく、双方とも空いている状態ならば共有記憶部Mを選択する。

⑤ 双方の通信手段について、全ての転送待ちデータ量の合計が同程度ならば通信回線Tの方が空いていると判断する。

⑤′ 双方の通信手段について、全ての転送待ちデータ量の合計が同程度ならば共有記憶部Mの方が空いていると判断する。

⑥ 待ち行列や通信待ちデータ量を比較する際、その差が予め定められた基準を越えていなければ通信回線Tの方が空いていると判断する。

⑥′ 待ち行列や通信待ちデータ量を比較する際、その差が予め定められた基準を越えていなければ共有記憶部Mの方が空いていると判断する。

⑦ 最近の或る単位時間に着目した各通信手段の

利用実績の統計に基づき、将来的に利用が少ないと予想される通信手段を空いていると判断する。

⑧ 共有記憶部Mのメッセージ送受信に用いる領域中、空き領域が少なくなれば混んでいると看做す。

⑨ 待ち行列の長さと転送待ちデータ量の合計とをそれぞれ比較する。そして

(a) 双方とも少なければ、その少ない方を空いていると判断する。

(b) 待ち行列の長さは長いが転送待ちデータ量の合計が少ない場合には、通信回線Tの方が空いていると判断する。

(b)' 待ち行列の長さは長いが転送待ちデータ量の合計が少ない場合には、共有記憶部Mの方が空いていると判断する。

(c) 或る比較基準の幅を設けておき、⑧⑧' に示すように判断する。

⑩ 或る時間帯については、常に一方の通信手段を空いていると看做す。

このような判断処理は、応用システムの形態やシステムの稼働状況に応じて、上述した判断基準の中から選択的に用いて、或いは適宜組み合わせて行われる。またこれらの判断基準を変数のパラメータとして表現しておき、オペレータ等からの指定により動的に上記パラメータを変更してその判断基準を変更するようにしても良い。

更には上述した判断基準によりその判断がつかない場合には、いずれかの通信手段を強制的に空いていると看做すようにしても良い。

その他、通信手段の選択戦略としては、常に負荷の少ない方を選択する場合の他に、一方の負荷をある範囲内に抑えるようにその選択制御を行うようにしても良い。このような手法はリアルタイム処理の発生が予想される場合等に有効である。例えばリアルタイムデータ参照要求に対して負荷の制約を課した方の通信手段を割り当てることにより、或る一定時間内にその処理を実行し得ることを概ね保証することが可能となる。この場合には前述した第5図に示すステップhの処理にて負荷を調整するように通信手段を割り当てることにより実現できる。

以上のような通信負荷状態の判断処理は、動的にその負荷状態を求める例であるが、通信の速度に基づきその利用方法を固定しておくことも可能である。例えば通信回線Tを利用した場合より共有記憶部Mを利用した場合の方が１００倍速く転送できることが予め分かっているような場合、１００回の通信要求のうち１回だけを通信回線Tに割り当て、残りの通信は共有記憶部Mを用いるようにすれば良い。

このような制御を行う場合であっても、前述した通信回線管理部7や通信回線負荷検出手段9，および共有記憶管理部8や共有記憶負荷検出手段10を用いることで、或いはその機能を簡単に変更するだけで容易に実現することができる。

ところで前記共有記憶部Mへのメッセージの書き込みは次のようにして行われる。基本的に前記共有記憶部Mへのメッセージの書き込みは、そのメッセージを受け取るべき計算機サイトの番号等の識別子を付加して行われる。各計算機サイトはこのような識別子を手掛かりとして前記共有記憶部Mに自己宛のメッセージが存在するか否か定期的に参照し、そのメッセージを受け取ることになる。或いは計算機サイトが共有記憶部Mにメッセージを書き込んだとき、そのメッセージを読み取るべき計算機サイトに対して割り込みをかけることで、メッセージが届いている旨を連絡するようにすることも可能である。

しかしてこの共有記憶部Mにおけるデータ管理には種々の手法がある。例えば共有記憶部Mに可変長で順次空いた領域にメッセージを書き込む手法や、共有記憶部Mの記憶領域を固定長のブロックに区切っておき、これらのブロックに順次メッセージを書き込んでいく手法がある。またメッセージを読み取ったか否かを判定する手段としては、読み取りが完了したメッセージを逐次消去していく手法や、そのメッセージが使用前であるか、或いは使用済みであるかを表す印（マーク）を付けておき、このマークを調べることで判定する手法等がある。

第６図はメッセージの受け取り先が複数ある場合、例えばデータのコピーが複数の計算機サイトにあって、それらの存在する全ての計算機サイトにデータロック，コミット等のメッセージを送信する場合の処理の流れを示す図である。また計算機サイトＡ，Ｂ，Ｃにそれぞれデータｘのコピーが存在し、これらの全てに対しロックを掛けると云う要求を２サイトから発した場合、前記共有記憶部Ｍに書き込まれるメッセージは、例えば第７図に示すように表現される。

この共有記憶部Ｍへのメッセージの書き込み処理について第６図を参照して説明すると、共有記憶管理部８は、メッセージを送るべき計算機サイトの数（参照カウンタ）の情報と共にそのメッセージを共有記憶部Ｍに書き込む（ステップｓ）。そして、例えば共有記憶部Ｍにメッセージを書き込んだことを、そのメッセージを受け取るべき計算機サイトに割り込みを掛けて通知する方式を採用した場合には、共有記憶管理部８はメッセージを届けた旨とそのメッセージの記述されている共有記憶部Ｍ上の番地をその通信先の計算機サイトに通知する（ステップｔ）。そしてこの通知を受けた計算機サイトでは、その共有記憶管理部８にてメッセージを読み取り入力し、前記参照カウンタをデクリメントする（ステップｕ）。

しかしてメッセージの送信側の計算機サイトの共有記憶管理部８では、前記参照カウンタを見ることでメッセージの読み取りを終了したサイトの数を検査し（ステップｖ）、参照カウンタが［０］になるまでメッセージの送信通知を繰り返す。そして前記参照カウンタが［０］になったら、そのメッセージが全ての送信先の計算機サイトで読み取られたことが確認されるので、その時点で前記メッセージを無効化する（ステップｗ）。尚、参照カウンタが［０］になった時点で自動的にそのメッセージを無効化するように予め規定しておくようにしても良い。

以上のようにして共有記憶部Ｍに対するメッセージの書き込みが制御され、共有記憶部Ｍを介する計算機サイト間でのメッセージの受け渡しが行われることになる。

ところで計算機サイトにおいて発生した命令がメッセージ受信要求である場合には、次のようにしてその処理手続きが進められる。このメッセージ受信要求は、例えば前述した第４図の［５］の処理に示されるようにデータの読み出しを要求した場合等、［６］の処理にて受け取るべきメッセージを受信要求したことになる。

しかしてメッセージ受信要求であることが検出された場合（ステップℓ）、先ず前記分散データベース通信制御部６は通信回線管理部７，或いは共有記憶管理部８のどちらにメッセージが到着したかを調べる（ステップｍ）。尚、前記共有記憶部Ｍを利用する場合には、ここでいう到着とはメッセージの実体そのものが共有記憶管理部８に到着した場合であってもよいが、メッセージが共有記憶部Ｍに書き込まれたことを共有記憶管理部８が確認した状態であっても良い。

このとき前記通信回線管理部７または共有記憶管理部８が前記分散データベース通信制御部６に割り込みをかけてメッセージ到着を知らせるように構成される場合と、そうでない場合とがある。後者の場合には通信回線管理部７と共有記憶管理部10のどちらにメッセージが到着しているか不明であることから、例えば分散データベース通信制御部６にて定期的に所定の間隔で前記通信回線管理部７と共有記憶管理部８とに対してメッセージが到着しているか否かを調べるようにすれば良い。このようなメッセージの受信問い合わせに対処するべく、共有記憶管理部８には共有記憶部Ｍのメッセージ領域を定期的に検査し、常にどのようなメッセージが共有記憶部Ｍに届いているかを管理する機能が組み込まれる。

また前記通信回線Ｔは分散データベース処理以外の処理にも使用される。そこで通信回線Ｔを介して通信されるメッセージ中には、分散データベース処理の為のメッセージであることを識別する情報が含められる。通信回線管理部７はこのようなメッセージを識別する為の情報を解釈し、分散データベース処理に関するメッセージであれば分

散データベース通信制御部8にメッセージが到着したことを割り込みを掛けて通知する。

分散データベース通信制御部8はこのようにしてメッセージの到着が確認される管理部から、そこに到着しているメッセージを受け取る（ステップn）。尚、分散データベース通信制御部6が共有記憶管理部8からメッセージを受け取るに際して、共有記憶部Mのアドレスが与えられるような場合には、共有記憶部Mからそのアドレスにあるメッセージを読み込むことで、そのメッセージの受信が行われる。

尚、計算機サイトが上述したメッセージの送受信処理以外の命令を実行する場合には、ステップoにおいて、例えば計算処理や外部記憶装置との入出力等に関する処理を行うことになる。

以上のようにしてこの実施例に係る分散データベースシステムでは、通信回線Tを介してメッセージ通信することのみならず、複数の計算機サイトにて共有される共有記憶部Mを用いて各計算機サイト間でのメッセージの受け渡しを行うものとなっている。

ところで前記共有記憶部Mから受け取るメッセージは、共有記憶部Mのデータ入出力速度が早いことから、通信回線Tから受け取るメッセージより高速に到達することがある。この為、どの通信手段を通してメッセージの通信が行われたかによって受信した複数のメッセージの順序が逆転することも考えられる。このような状態を検出して受信メッセージを正しい順序で処理を行う為には、例えばメッセージにその順序を示す為の番号等を付加しておくようにすれば良い。そして、例えばメッセージ再構成手段等を設けておくことにより、上記メッセージ番号に基づいて受信メッセージをそのメッセージ順序に従って処理するようにすれば良い。またメッセージ単位ではなく、データ参照要求を単位として通信手段を決定するような手法を採用すれば、メッセージの順序は常に正しく保たれることになる。

以上のようにして通信回線Tと共有記憶部Mとを用いてメッセージ通信を行うようにした実施例によれば、その通信効率を飛躍的に高めることができる。

この通信効率の改善効果について説明すると、前述した第４図に示す通信処理手続きの例では計算機サイト間でのメッセージ通信回数は１８回である。これをサイト間の通信という観点から図示すると第８図に示すようになる。尚、第８図中の番号は、第４図におけるメッセージの番号を示している。

ここで本発明の効果を示すべく、或る仮定のもとに通信所要時間のみに着目して概算してみると次のようになる。

例えば通信回線Tに比較して共有記憶部Mを用いた方が１００倍高速に通信できるものとする。即ち、通信回線Tによる通信時間，共有記憶部Mに書き込む時間と共有記憶部Mから読み出す時間との和の比が［１００：１］であると仮定する。

またデータベースを構成するデータの通信量が、その制御情報に比較して１０倍多いと仮定する。

このような仮定の下で共有記憶部Mを用いた場合の制御情報に関するメッセージ通信を１単位時間とするとき，従来の方式において必要とする通信所要時間は次のようになる。

ＡＢ間：制御に関するメッセージ
（７回×１００）単位時間
データに関するメッセージ
（２回×１０×１００）単位時間

ＡＣ間：制御に関するメッセージ
（７回×１００）単位時間
データに関するメッセージ
（２回×１０×１００）単位時間

合　計　（５，４００）単位時間

これに対して本発明を採用した場合、例えばＡ－Ｂ間の通信回数のうち約１０％，Ａ－Ｃ間の通信回数のうち約２０％を通信回線Ｔを用いて行い、それ以外については共有記憶部Ｍを用いて通信したと仮定すると、つまりデータの通信［６］，［８］，［９］，［１０］を共有記憶部Ｍを用いて通信したと仮定すると次のようになる。

ＡＢ間：通信回線Ｔを使用するメッセージ
（１回×１００）単位時間
共有記憶部Ｍを使用するメッセージ中
＊制御に関するメッセージ
（６回×１）単位時間
＊データに関するメッセージ
（２回×１０）単位時間
ＡＣ間：通信回線Ｔを使用するメッセージ
（２回×１００）単位時間
共有記憶部Ｍを使用するメッセージ中
＊制御に関するメッセージ
（５回×１）単位時間
＊データに関するメッセージ
（２回×１０）単位時間
合　計　（３５１）単位時間

このような通信所要時間の対比結果から明らかなように、通信処理の部分に焦点を当てればメッセージ通信について約１５倍の高速化が実現できる。従って本発明による効果が非常に大きいことが分かる。また現在はＣＰＵ処理やディスクの入出力以上に通信に多くの時間を要していることから、上述した本発明によって分散データベース参照の飛躍的な高速処理を実現し得ることが期待できる。

またここで取り上げた例では、データベース内にデータの重複（コピーという）を考慮していないが、データの安全性・読み出しの高速化を図るべく、そのコピーを複数サイトに配置する場合は、ロックやコミットメントに関するメッセージが必然的に多くなる。従って、複数の計算機サイトに多くのコピーが存在する分散データベースでは特にその効果が大きくなる。

尚、本発明は上述した実施例に限定されるものではなく、計算機サイト内の処理部の構成方法等は、その要旨を逸脱しない範囲で適宜変形して実現することができる。例えば第１図では共有記憶部Ｍは全ての計算機サイトに結合され直接参照可能となっているが、特定の計算機サイトにのみ共有記憶部Ｍが接続されている場合であっても、或る程度本発明の効果が得られる。更に本発明における通信制御の形態は、質問最適化方式，並行制御方式，コミットメント方式等のデータベース制御とは全く独立したものであるから、それぞれについてどのようなアルゴリズムを採用した場合にも適用可能である。その他、本発明はその要旨を逸脱しない範囲で種々変形して実施することができる。

［発明の効果］

以上説明したように本発明によれば、複数の計算機サイトを相互に結合する通信回線Ｔに加えて、上記各計算機サイトに共有される共有記憶手段を備え、分散データベース処理において数多く発生するメッセージを、その負荷状態に応じて通信回線と共有記憶手段とを使い分けて通信するので、分散データベースへの参照要求を高速に処理することが可能となる。

また本発明によればシステム全体のスループットを増大させるだけではなく、従来、通信ネックにより実現できなかった処理や、データ参照要求やメッセージに優先順位をつける必要がある場合等、その処理が難しかったリアルタイム処理と通常の処理の混在したデータ参照処理も簡単に実現できる。例えば計算機サイト間で大容量のデータ同志を結合するような場合には一般的には通信回線Ｔをデータ転送で占有してしまうことになる。そしてリアルタイム性を要求されるデータ参照処理に支障をきたすことになる。このような場合であっても通信回線Ｔまたは共有記憶部Ｍの一方の負荷をある決められた範囲内に抑えておくような戦略を採用することにより、リアルタイムなデータ参照処理も高速に実現することが可能となる。

更には通信回線Ｔまたは共有記憶部Ｍの一方が故障したときでも、例えば通信回線負荷検出手段9の負荷を無限大にすることで他方で代替でき、その処理速度の低下は懸念されるがデータ参照要求の処理を継続することが可能となる。この結果、システムの信頼性を向上させることができるという効果が奏せられる。

４．図面の簡単な説明

第１図は本発明の一実施例に係る分散データ

ベース処理方式を採用した分散データベースシステムの全体構成を示すブロック図、第２図は実施例システムにおける計算機サイトの構成例を示す図、第３図は分散データベース参照要求の例に基づくデータ参照処理手順を説明するための図、第４図はデータ参照要求を処理するときに発生する通信すべきメッセージの例を示す図である。

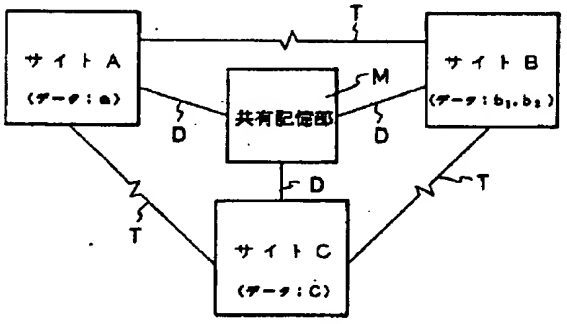
また第５図は分散データベース参照の為の高速通信処理を実現したときの通信部分に焦点を当てた処理の流れを示す図、第６図は共有記憶部を利用して複数サイトに同一のメッセージを送信する場合の共有記憶部の管理手順を示す図、第７図は共有記憶を利用して複数サイトに同一のメッセージを送信する場合の共有記憶部上でのメッセージの構成例を示す図、第８図はデータ参照要求処理するときのサイト間通信メッセージを説明する為の図である。

そして第９図は従来の分散データベースシステムの構成例を示す図、第１０図は従来システムにおける計算機サイトの構成例を示す図である。

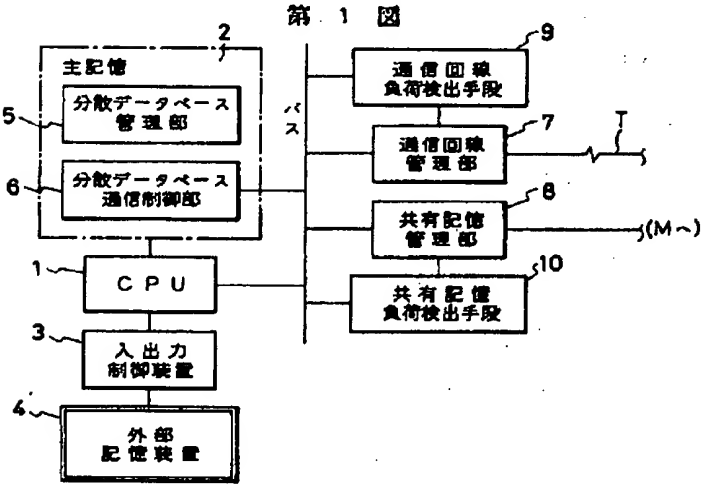
Ａ，Ｂ，Ｃ…計算機サイト、Ｔ…通信回線、Ｍ…共有記憶部、Ｄ…情報伝送路、

１…ＣＰＵ、２…主記憶、３…入出力制御部、４…外部記憶装置、５…分散データベース管理部、６…分散データベース通信制御部、７…通信回線管理部、８……共有記憶管理部、９…通信回線負荷検出手段、１０…共有記憶負荷検出手段。

出願人代理人　弁理士　鈴江武彦




第１図



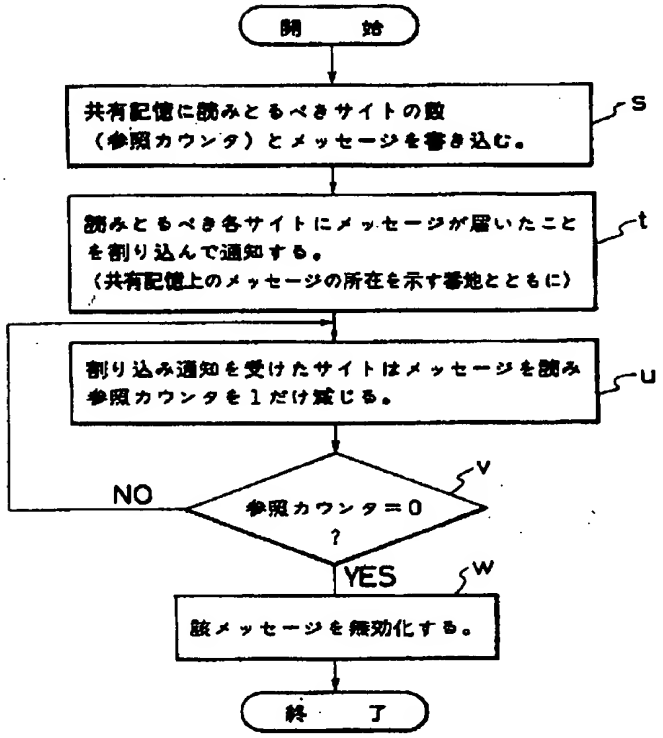

第２図

```
Begin-transaction
    read        x, B:b1
      /*  変数xにBサイトのb1を読み込む.   */
    read        y, A:a
      /*  変数yにAサイトのaを読み込む.    */
    read        z, C:c
      /*  変数zにCサイトのcを読み込む.    */
    z = z - ( x - y )
      /*  各サイトから読んできたデータに基づき
          値を計算する.                   */
    write       C:c, z
      /*  変数zをCサイトのcに書き込む.    */
    write       B:b2, z
      /*  変数zをBサイトのb2に書き込む.   */
End-transaction
```

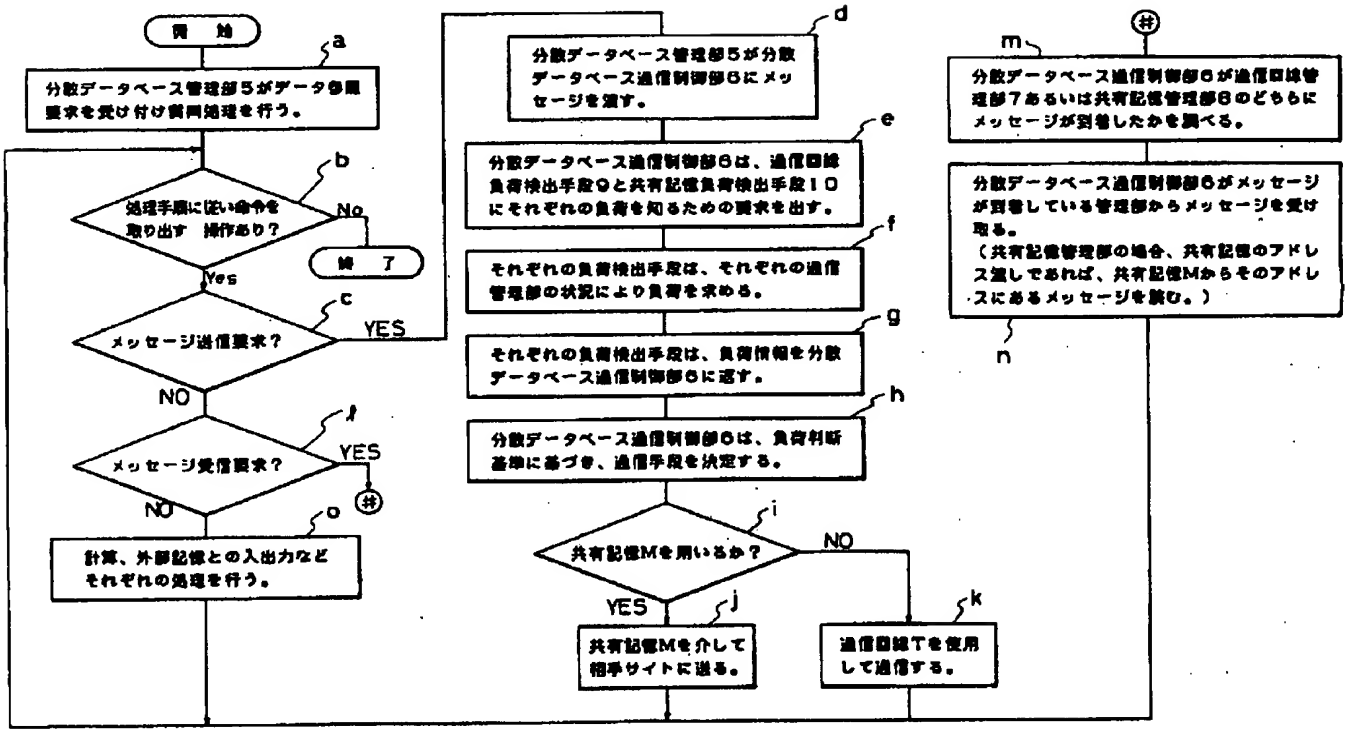
第３図

| メッセージ | 送信サイト | 受信サイト |
|---|---|---|
| (1) LOCK (b1, b2) | A | B |
| (2) LOCK-END | B | A |
| (3) LOCK (c) | A | C |
| (4) LOCK-END | C | A |
| (5) READ (b1, b2) | A | B |
| (6) DATA (b1, b2) | B | A |
| (7) READ (c) | A | C |
| (8) DATA (c) | C | A |
| (9) WRITE (c) | A | C |
| (10) WRITE (c) | A | B |
| (11) COMMIT1 | A | B |
| (12) COMMIT-OK | B | A |
| (13) COMMIT1 | A | C |
| (14) COMMIT-OK | C | A |
| (15) COMMIT2 | A | B |
| (16) COMMIT-OK | B | A |
| (17) COMMIT2 | A | C |
| (18) COMMIT-OK | C | A |

第 4 図




第 6 図

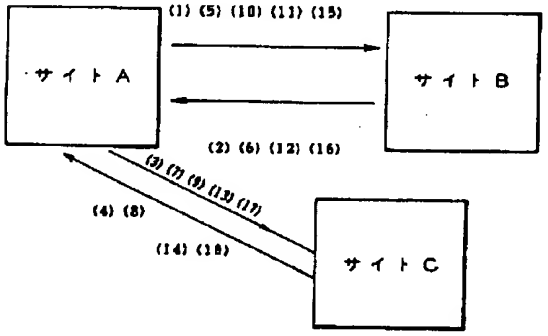



第 5 図

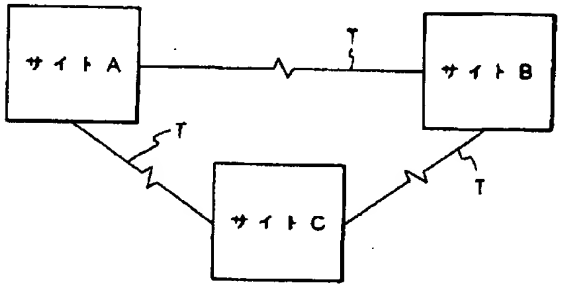
| メッセージ識別子 | メッセージサイズ (BYTE) | 読み取りサイト数 (参照カウンタ) | 発信サイト名 | 受信サイト名 |
|---|---|---|---|---|
| ××××××××× | 256 | 3 | Z | A, B, C |

| データ参照要求番号 | メッセージ内容 |
|---|---|
| ソソソソソソソソ | LOCK(x) |

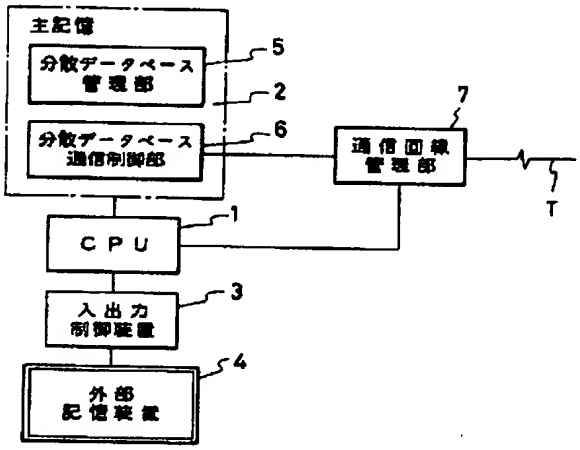
第 7 図




第 8 図




第 9 図




第 10 図

Reference 1: JP 03-198133 A

FIG.1

サイトA: SITE A (DATA: A)

サイトB: SITE B (DATA: $B_1$, $B_2$)

サイトC: SITE C (DATA: C)

共有記憶部: SHARED STORAGE SECTION

FIG.2

1: CPU

2: MAIN MEMORY

3: I/O CONTROLLER

4: EXTERNAL MEMORY

5: DISTRIBUTED DATABASE MANAGING SECTION

6: DISTRIBUTED DATABASE COMMUNICATION CONTROLLER

7: COMMUNICATION LINE MANAGING SECTION

8: SHARED STORAGE MANAGING SECTION

9: COMMUNICATION LINE LOAD DETECTING SECTION

10: SHARED STORAGE LOAD DETECTING SECTION

バス: BUS

[line 12 in upper right column of page 4 to line 8 in lower right column of page 4]

A distributed database processing system according to one embodiment of the present invention will be described below with reference to the accompanying drawings.

FIG.1 is a schematic block diagram showing a distributed database processing system configured using the preferred embodiment system. Here, an

embodiment of the system including three computer sites A, B and C is shown. The respective computer sites A, B and C are interconnected via a communication line T. Further, a shared storage section M shared with the respective computer sites A, B and C is provided. This shared storage section M is a section including a semiconductor memory and configured as follows. That is, the section M is connected to the respective computer sites A, B and C via an information transmission channel D such as a bus so as to enable a data reference from each of the computer sites A, B and C at high speeds.

Thus, the respective computer sites A, B and C are interconnected via a communication line T and at the same time, share the shared storage section M. Schematically, the computer site comprises, as a section for controlling communication with the other communication sites, a communication line managing section for managing the communication using the communication line T and a shared storage managing section for managing the data reference to the shared storage section M. In addition, the computer site is configured as follows. That is, the computer detects the congestion in the communication line T and the shared storage section M using, for example, a communication line load detecting section and a shared storage load detecting section, respectively. Then, based on these detection results of the loading state, the computer transmits a message communication request to the communication line managing section or the shared storage managing section to perform data transmission and reception to and from the other computer sites selectively via the communication line T or the shared storage section M.

In short, the computer site is configured, for example, as shown in FIG.2. This computer site has a function for performing processing as a computer by itself and is mainly composed of a CPU 1 and a main memory 2. To the CPU 1, an external memory 4 such as a magnetic disk device for storing a database is connected via an I/O controller 3 such as an I/O channel for managing an input-output of the CPU 1.